

PATENT APPLICATION

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re the Application of:

LIUHTO et al

Group Art Unit: Not yet assigned

Application No.: New Continuation Appln.

Examiner: Not yet assigned

Filed: July 9, 2003

Attorney Dkt. No.: 60279.00057

For: SYNCHRONIZING LOCAL NETWORK

CLAIM FOR PRIORITY UNDER 35 USC § 119

Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

July 9, 2003

Sir:

The benefit of the filing dates of the following prior foreign application(s) filed in the following foreign country(ies) is hereby requested for the above-identified patent application and the priority provided in 35 U.S.C. §119 is hereby claimed:

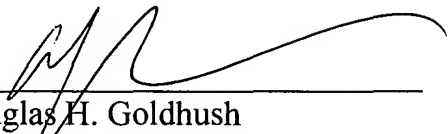
Finnish Patent Application No. 20012173 filed on November 9, 2001 in Finland

In support of this claim, certified copy(ies) of said original foreign application(s) is filed herewith.

It is requested that the file of this application be marked to indicate that the requirements of 35 U.S.C. §119 have been fulfilled and that the Patent and Trademark Office kindly acknowledge receipt of this document(s).

Please charge any fee deficiency or credit any overpayment with respect to this paper to Counsel's Deposit Account No. 50-2222.

Respectfully submitted,



Douglas H. Goldhush
Registration No. 33,125

Customer No. 32294
SQUIRE, SANDERS & DEMPSEY LLP
14TH Floor
8000 Towers Crescent Drive
Tysons Corner, Virginia 22182-2700
Telephone: 703-720-7800
Fax: 703-720-7802

DHG:scc

Enclosure: Priority Document(s) (1)

PATENTTI- JA REKISTERIHALLITUS
NATIONAL BOARD OF PATENTS AND REGISTRATION

Helsinki 13.5.2003

ETUOIKEUSTODISTUS
PRIORITY DOCUMENT



Hakija
Applicant **Nokia Corporation**
Helsinki

Patenttihakemus nro
Patent application no **20012173**

Tekemispäivä
Filing date **09.11.2001**

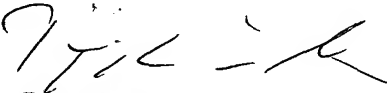
Kansainvälinen luokka
International class **H04L**

Keksinnön nimitys
Title of invention

"Synchronizing local network"
(Synkronisoiva paikallisverkko)

Täten todistetaan, että oheiset asiakirjat ovat tarkkoja jäljennöksiä Patentti- ja rekisterihallitukselle alkuaan annetuista selityksestä, patenttivaatimuksista, tiivistelmästä ja piirustuksista.

This is to certify that the annexed documents are true copies of the description, claims, abstract and drawings originally filed with the Finnish Patent Office.


Pirjo Kaila
Tutkimussihteeri

Maksu 50 €
Fee 50 EUR

Maksu perustuu kauppa- ja teollisuusministeriön antamaan asetukseen 1027/2001 Patentti- ja rekisterihallituksen maksullisista suoritteista muutoksineen.

The fee is based on the Decree with amendments of the Ministry of Trade and Industry No. 1027/2001 concerning the chargeable services of the National Board of Patents and Registration of Finland.

Osoite:	Arkadiankatu 6 A	Puhelin:	09 6939 500	Telefax:	09 6939 5328
	P.O.Box 1160	Telephone:	+ 358 9 6939 500	Telefax:	+ 358 9 6939 5328
	FIN-00101 Helsinki, FINLAND				

SYNCHRONIZING LOCAL NETWORK

FIELD OF THE INVENTION

5 The present invention relates to telecommuni-
cations. In particular, the present invention relates
to a novel and improved system for transmitting inter-
nal messages in a local network.

BACKGROUND OF THE INVENTION

10 Fault tolerant local computer networks are
used when high reliability is required. Typical exam-
ples are modern telecommunication networks where sys-
tem downtime must be minimized in order to provide
service of acceptable quality. A common way to imple-
15 ment fault tolerance is to use replicated elements in
the networks. A network with replicated application
processes is a network where each application process
is simultaneously running in at least two separate
computer units. Thus, if a computer unit running a
20 given application process goes down for whatever rea-
son, there's still at least one more computer unit
running the same process.

 Message synchronism is one possible way to
implement a network with replicated elements. In other
25 words, when a transmitting application process sends a
message to a receiving application process, an identi-
cal copy of the message is sent to each replicated (or
redundant) receiving process running in various com-
puter units. Each replicating receiving process must
30 receive an identical message and in identical order to
maintain synchronization of messages. Message synchro-
nism is sometimes also referred to as input synchro-
nism.

 A network with message synchronism can be im-
35 plemented either by software or by hardware. When im-
plemented by software, a special software layer is im-

plemented on top of a standard messaging protocols or technologies such as Ethernet or IP (Internet Protocol, IP). However, software based messaging consumes a substantial amount of computing power and network bandwidth.

Another way to implement a network which takes care of the message synchronism is to use special proprietary hardware, typically a message bus. When a message is sent from one computer unit to another, it is replicated in the message bus hardware. The message bus acts as a serializing point and the message order is preserved.

A typical example of such a hardware based prior art implementation is illustrated in figure 1. The system comprises multiple sending computer units CPUs, each for running at least one sending application process for sending an internal message. Since the system in question implements fault tolerance with replicated elements, each message is sent to two or more recipients. The system further comprises multiple receiving computer units CPUr, each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units.

The system further comprises one proprietary interface unit IF per one or more computer units for buffering and relaying messages sent to and from the corresponding computer unit. Each interface unit comprises a transmitting buffer TX for storing one or more message to be sent until processed by the internal bus, and a receiving buffer RX for storing one or more received messages until processed by the corresponding computer unit.

The system further comprises multiple internal links, such as e.g. Compact PCI -buses (Peripheral

Component Interconnect, PCI), each for linking a computer unit to its corresponding interface unit.

The system further comprises a proprietary external message bus for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time. The external message bus is linked with the interface units. The external message bus is a shared bus, i.e. only one sent message is forwarded by the bus at any given time, thus message synchronization is maintained.

However, there are significant disadvantages with the referenced prior art implementation. When more computing capacity is needed, more computer units are attached to the message bus. Thus a longer message bus is needed. Adding more computer units requires more messaging capacity between the computer units. The physical length of the message bus is limited because it is a shared bus. There is a trade-off between the length of the bus and the speed (bits per second). The higher the speed, the shorter the bus must be. The speed of the bus can be made higher by using more bits in parallel but the physical limits such as connectors and cabling will be reached very soon. Typically, the existing implementations use 8 or 16 bits in parallel.

Thus there is need for a solution overcoming the capacity limitations of the present message bus based implementations of a fault tolerant, high capacity synchronized networks.

Because the functionalities required from the system for providing message synchronism are in direct violation of present related specifications, commercially available solutions, e.g. Ethernet switches are not suitable. This stems from the fact that the specification IEEE802.1D prohibits the switch sending a

message back to the same computer unit that the message was received from. Yet such a functionality is required in order to provide message synchronism in all situations. Thus current implementations, such as
5 the one disclosed in fig. 1, need to use proprietary components in order to achieve the desired functionalities.

Thus there is need for a solution making it possible to implement a system providing message synchronism using non-proprietary components either as-is
10 or with only slight modifications.

SUMMARY OF THE INVENTION

The present invention concerns a system for
15 transmitting internal messages in a local network while maintaining message synchronism. An internal message refers to a message sent from a computer unit in a local network to computer units in the same local network as opposed to traffic to/from external networks such as e.g. Internet. The system comprises multiple sending computer units, each for running at least one sending application process for sending an internal message. The system further comprises multiple receiving computer units, each for running at least one receiving application process for receiving
20 a sent internal message, at least two copies of each receiving application process residing in said receiving computer units. The term receiving application process refers to application processes that for whatever reason are interested in receiving a sent message in identical order. In an embodiment of the present invention the receiving application processes are an application process and its replicated copies. The division of computer units into sending and receiving
25 ones is a functional one, thus a computer unit can simultaneously be both sending and receiving messages. Correspondingly, an application process can simultane-

ously be both sending and receiving messages. The computer units are physically or logically separate from each other.

According to the invention, the system further comprises one interface unit per one or more computer units for buffering and relaying messages sent to and from the corresponding computer units.

Further according to the invention, the system comprises multiple external links, each for linking a computer unit to its corresponding interface unit. In an embodiment of the invention said external links are fast serial links.

Further according to the invention, the system comprises an internal interconnecting device for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining message synchronism. The interconnecting device is internally coupled with the interface units.

Thus also messages sent and received by application processes running in the same computer unit are routed via the interconnecting device. Since the long and thus relatively slow external message bus of prior art is replaced with a short and extremely fast internal interconnecting device, the overall messaging capacity of the system increases while message synchronization is maintained. Yet, because of external links used between computer units and interface units, distance between the computer units can still be relatively long, i.e. tens of meters.

Another embodiment of the invention concerns a system for transmitting internal messages in a local network while maintaining message synchronism. The system comprises multiple sending computer units, each for running at least one sending application process

for sending an internal message. The system further comprises multiple receiving computer units, each for running at least one receiving application process for receiving a sent internal message, at least two copies
5 of each receiving application process residing in said receiving computer units. The computer units are physically or logically separate from each other.

According to this embodiment of the invention, the system comprises multiple multiplexer units,
10 each for collecting messages from and distributing messages to one or more sending computer units.

Further according to this embodiment of the invention, the system further comprises one interface unit per one or more multiplexer units for buffering
15 and relaying messages sent to and from the corresponding multiplexer units. In an embodiment of the invention there may be multiple layers of multiplexer units, thus a multiplexer unit may be connected to the interface unit via another multiplexer unit. By doing
20 so the number of computer units linked to the internal interconnecting device using one interface unit can be increased. At the same time the amount of cabling may be reduced.

Further according to this embodiment of the invention, the system comprises an internal interconnecting device for receiving messages relayed by the
25 interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining message synchronism. The interconnecting device is internally coupled with the interface units. Because of multiplexers used between computer units and interface units, cabling is reduced
30 significantly.

In an embodiment of the invention each interface unit further comprises a transmitting buffer for

storing one or more messages to be sent until processed by the interconnecting device, and a receiving buffer for storing one or more received messages until processed by the corresponding computer or multiplexer
5 unit.

In an embodiment of the invention messages are sent as multicasts by the sending application process. Multicasting can be done e.g. by using group addressing. Group addressing refers to a known addressing method where a group of computer units are
10 assigned a common group address. A message to be sent to computer units of a given group is addressed according to corresponding group address. With two recipients group addressing is commonly referred to as
15 pair addressing.

In an embodiment of the invention the interconnecting device is an internal bus.

In an embodiment of the invention the interconnecting device is a crossbar.

20 In an embodiment of the invention the interconnecting device, the interface units coupled to it and the multiplexer units are implemented by modifying a standard LAN switch such as an Ethernet switch. Such modifications comprise modifications making it possible
25 to multicast messages back to the same computer unit that the message was received from, as well as modifications required to make the interconnecting device act as a serializing point.

The invention makes it possible to overcome
30 the capacity limitations of the present external message bus based proprietary implementations of a fault tolerant, high capacity synchronized networks. Significantly higher messaging capacity can be achieved than with existing implementations while at the same
35 time maintaining the message synchronization required by the fault tolerant, high capacity synchronized networks. Additionally, the invention makes it possible

to implement a system providing message synchronism using non-proprietary components either as-is or with only slight modifications. The system can be implemented e.g. by only slightly modifying already existing LAN switches.

BRIEF DESCRIPTION OF THE DRAWINGS

The accompanying drawings, which are included to provide a further understanding of the invention and constitute a part of this specification, illustrate embodiments of the invention and together with the description help to explain the principles of the invention. In the drawings:

Fig 1 is a block diagram illustrating a prior art system,

Fig 2 is a block diagram illustrating a system according to one embodiment of the present invention, and

Fig 3 is a block diagram illustrating a system according to another embodiment of the present invention.

DETAILED DESCRIPTION OF THE INVENTION

Reference will now be made in detail to the embodiments of the present invention, examples of which are illustrated in the accompanying drawings.

Figure 2 illustrates a system for transmitting internal messages in a local network while maintaining message synchronism. The system comprises multiple sending computer units CPUs (only one of which is disclosed in Figure 2), each for running at least one sending application process for sending an internal message. The system further comprises multiple receiving computer units CPU_r, each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiv-

ing application process residing in said receiving computer units. Application processes are elements performing information processing for a given application.

5 The system illustrated in Figure 2 further comprises one interface unit IF per one or more computer units for buffering and relaying messages sent to and from the corresponding computer units. In the embodiment of the invention illustrated in Figure 2
10 messages are sent as multicasts by the sending application process. Each interface unit IF further comprises a transmitting buffer TX for storing one or more messages to be sent until processed by the interconnecting device, and a receiving buffer RX for storing one or more received messages until processed by
15 the corresponding computer unit. Preferably the buffers are FIFO (First In - First Out) buffers.

 The system illustrated in Figure 2 further comprises multiple external links SrL, each for linking a computer unit to its corresponding interface
20 unit. In the embodiment of the invention illustrated in Figure 2 said external links are fast serial links.

 The system illustrated in Figure 2 further comprises an internal interconnecting device IxD for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining
25 message synchronism. The interconnecting device is internally coupled with the interface units. In the embodiment of the invention illustrated in Figure 2 said interconnecting device is an internal message bus. However, the interconnecting device can also be implemented e.g. as a crossbar or as some other internal
30 bus arrangement known to a person skilled in the art. Messages sent and received by application processes

running in the same computer unit are also routed via the internal bus.

Figure 3 illustrates a system for transmitting internal messages in a local network while maintaining message synchronism. The system comprises multiple sending computer units CPUs, each for running at least one sending application process for sending an internal message. The system further comprises multiple receiving computer units CPUr, each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units. Application processes are elements performing information processing for a given application.

The system illustrated in Figure 3 further comprises multiple multiplexer units MUX, each for collecting messages from and distributing messages to one or more sending computer units. The multiplexers collect traffic and messages from one or more computer units and deliver them to the internal interconnecting device via the interface units. The multiplexers are used to reduce cabling and the number of interfaces or ports in the interface units and the interconnecting device. The multiplexers may be implemented so that they pass all the traffic from computer units to interface units. Additionally the multiplexers can be implemented so that unicast messages sent to the computer units residing in the area of the same multiplexer are directed to the correct computer units. However in case of multicast messages, if any of the addressed computer units or application processes is not residing in the area of said multiplexer, the message is directed to the right addresses via the interface units and the interconnecting device. Otherwise the message synchronism can not be maintained. In the embodiment of the invention illustrated in Figure 3

there is one layer of multiplexer units. However, the invention can also be implemented with multiple layers of multiplexer units.

The system illustrated in Figure 3 further comprises one interface unit IF per one or more multiplexer units for buffering and relaying messages sent to and from the corresponding multiplexer units. In the embodiment of the invention illustrated in Figure 3 messages are sent as multicasts by the sending application process. Each interface unit IF further comprises a transmitting buffer TX for storing one or more messages to be sent until processed by the interconnecting device, and a receiving buffer RX for storing one or more received messages until processed by the corresponding multiplexer unit. Preferably the buffers are FIFO (First In - First Out) buffers.

The system illustrated in Figure 3 further comprises an internal interconnecting device IxD for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time thus maintaining message synchronism. The interconnecting device is internally coupled with the interface units. In the embodiment of the invention illustrated in Figure 3 said interconnecting device is an internal message bus. However, the interconnecting device can also be implemented e.g. as a crossbar or as some other internal bus arrangement known to a person skilled in the art. Messages sent and received by application processes running in the same computer unit are also routed via the internal bus.

It is obvious to a person skilled in the art that with the advancement of technology, the basic idea of the invention may be implemented in various ways. The invention and its embodiments are thus not

limited to the examples described above, instead they may vary within the scope of the claims.



CLAIMS

1. A system for transmitting internal messages in a local network while maintaining message synchronism, comprising:

5 multiple sending computer units (CPUs), each for running at least one sending application process for sending an internal message, said message being sent to two or more recipients, and

multiple receiving computer units (CPUr),
10 each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units,

characterized in, that the system
15 further comprises:

one interface unit (IF) per one or more computer units for buffering and relaying messages sent to and from the corresponding computer units,

multiple external links (SrL), each for linking
20 a computer unit to its corresponding interface unit, and

an internal interconnecting device (IxD) for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units
25 corresponding to the respective receiving computer units one received message at a time, said interconnecting device internally coupled with the interface units.

30 2. The system according to claim 1, characterized in that each interface unit further comprises:

a transmitting buffer (TX) for storing one or more message to be sent until processed by the inter-
35 connecting device, and

a receiving buffer (RX) for storing one or more received messages until processed by the corresponding computer unit.

3. The system according to claims 1 or 2, characterized in that messages are sent as multicasts by the sending application process.

4. The system according to claims 1, 2 or 3, characterized in that messages sent and received by application processes running in the same computer unit are routed via the interconnecting device.

5. The system according to claims 1, 2, 3 or 4, characterized in that the interconnecting device is an internal bus.

6. The system according to claims 1, 2, 3 or 4, characterized in that the interconnecting device is a crossbar.

7. The system according to claims 1, 2, 3, 4, 5 or 6, characterized in that the interconnecting device and the interface units coupled to it are implemented as a modified LAN switch.

8. A system for transmitting internal messages in a local network while maintaining message synchronism, comprising:

multiple sending computer units (CPUs), each for running at least one sending application process for sending an internal message, said message being sent to two or more recipients using group addressing, and

multiple receiving computer units (CPr), each for running at least one receiving application process for receiving a sent internal message, at least two copies of each receiving application process residing in said receiving computer units,

characterized in, that the system further comprises:

multiple multiplexer units (MUX), each for collecting messages from and distributing messages to one or more sending computer units,

5 one interface unit (IF) per one or more multiplexer units for buffering and relaying messages sent to and from the corresponding multiplexer units, and

10 an internal interconnecting device (IxD) for receiving messages relayed by the interface units corresponding to the sending computer units, and for forwarding each received message to the interface units corresponding to the respective receiving computer units one received message at a time, said interconnecting device internally coupled with the interface
15 units.

9. The system according to claim 8, characterized in that each interface unit further comprises:

20 a transmitting buffer (TX) for storing one or more message to be sent until processed by the interconnecting device, and

a receiving buffer (RX) for storing one or more received messages until processed by the corresponding multiplexer unit.

25 10. The system according to claims 8 or 9, characterized in that messages are sent as multicasts by the sending application process.

30 11. The system according to claims 8, 9 or 10, characterized in that messages sent and received by application processes running in the same computer unit are routed via the interconnecting device.

35 12. The system according to claims 8, 9, 10 or 11, characterized in that the interconnecting device is an internal bus.

13. The system according to claims 8, 9, 10 or 11, characterized in that the interconnecting device is a crossbar.

5 14. The system according to claims 8, 9, 10, 11, 12 or 13, characterized in that the interconnecting device, the interface units coupled to it and the multiplexer units are implemented as modified LAN switches.

10 15. The system according to claims 8, 9, 10, 11, 12, 13 or 14, characterized in that a multiplexer unit is connected to an interface unit via another multiplexer unit.

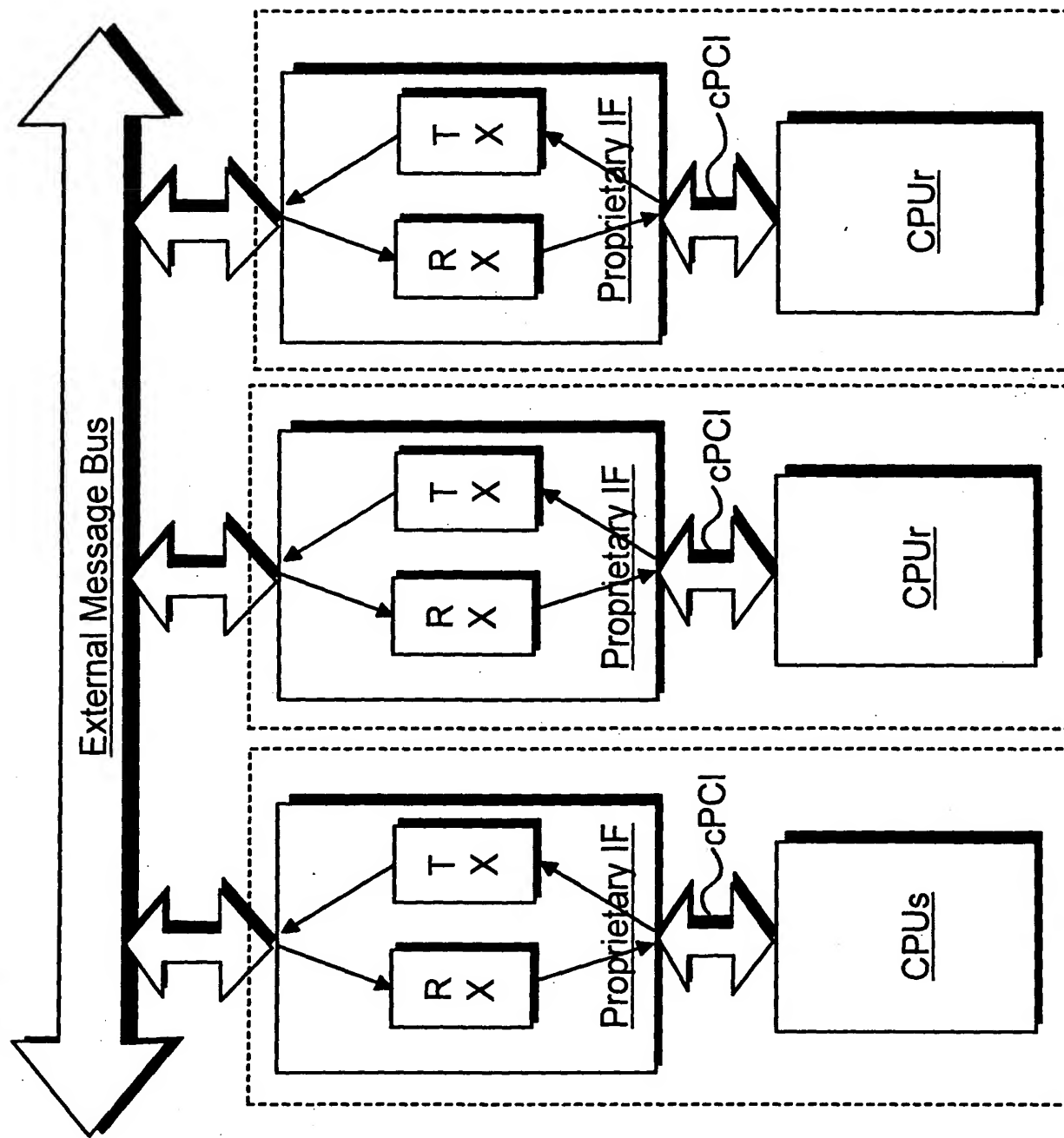


(57) ABSTRACT

The present invention concerns a system for transmitting internal messages in a synchronizing local network. The system comprises multiple computer units for running application processes. The system further comprises multiple interface units, each for transmitting an internal message from a sending computer unit to several receiving computer units. According to the invention, the system further comprises multiple external links, each for linking a computer unit to an interface unit. The system further comprises an interconnecting device for receiving messages from sending computer units and forwarding them to receiving computer units one sent message at a time.

(Fig. 2)

2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100



Prior Art.

Fig. 1

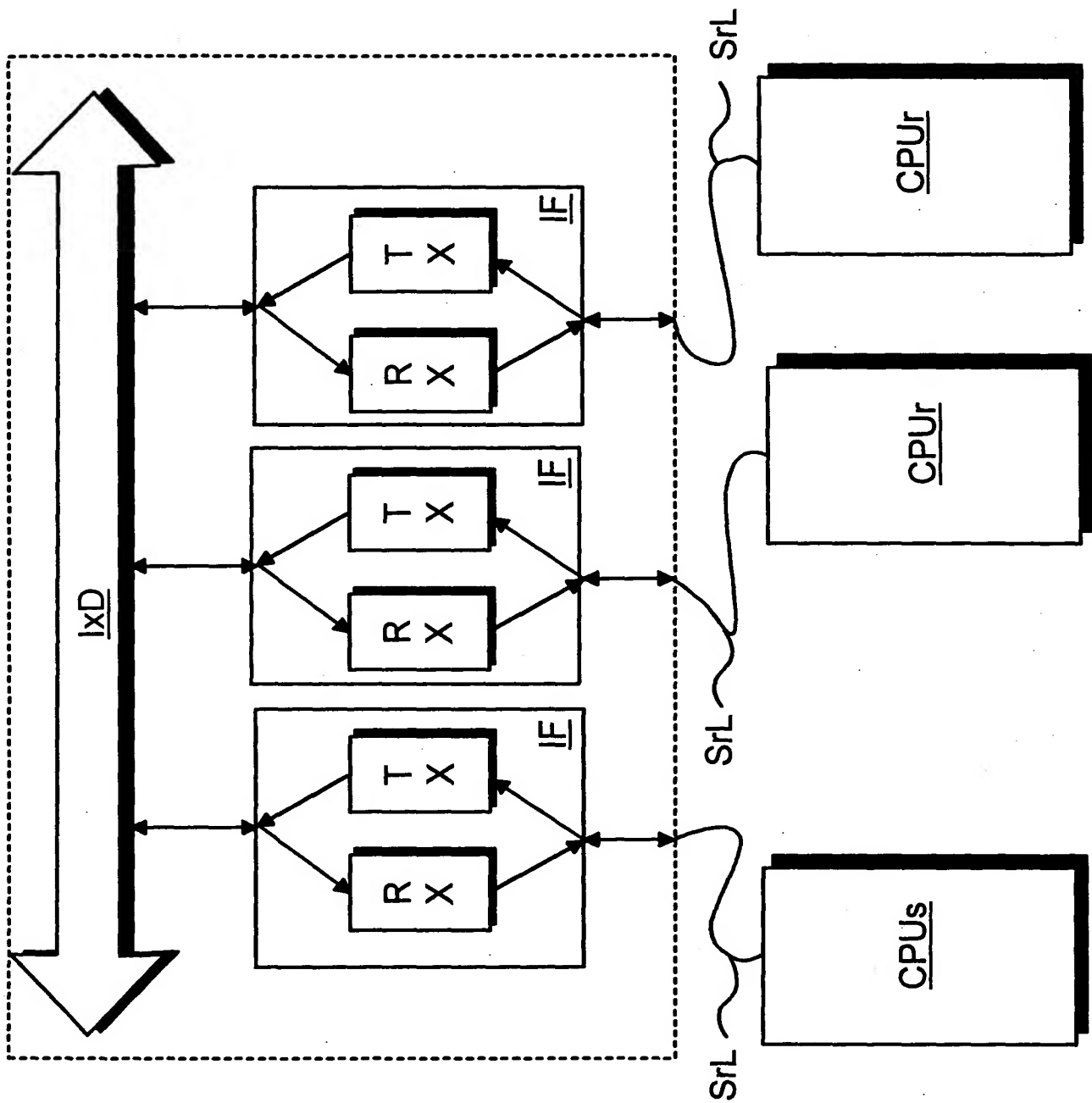


Fig. 2

001101 010100

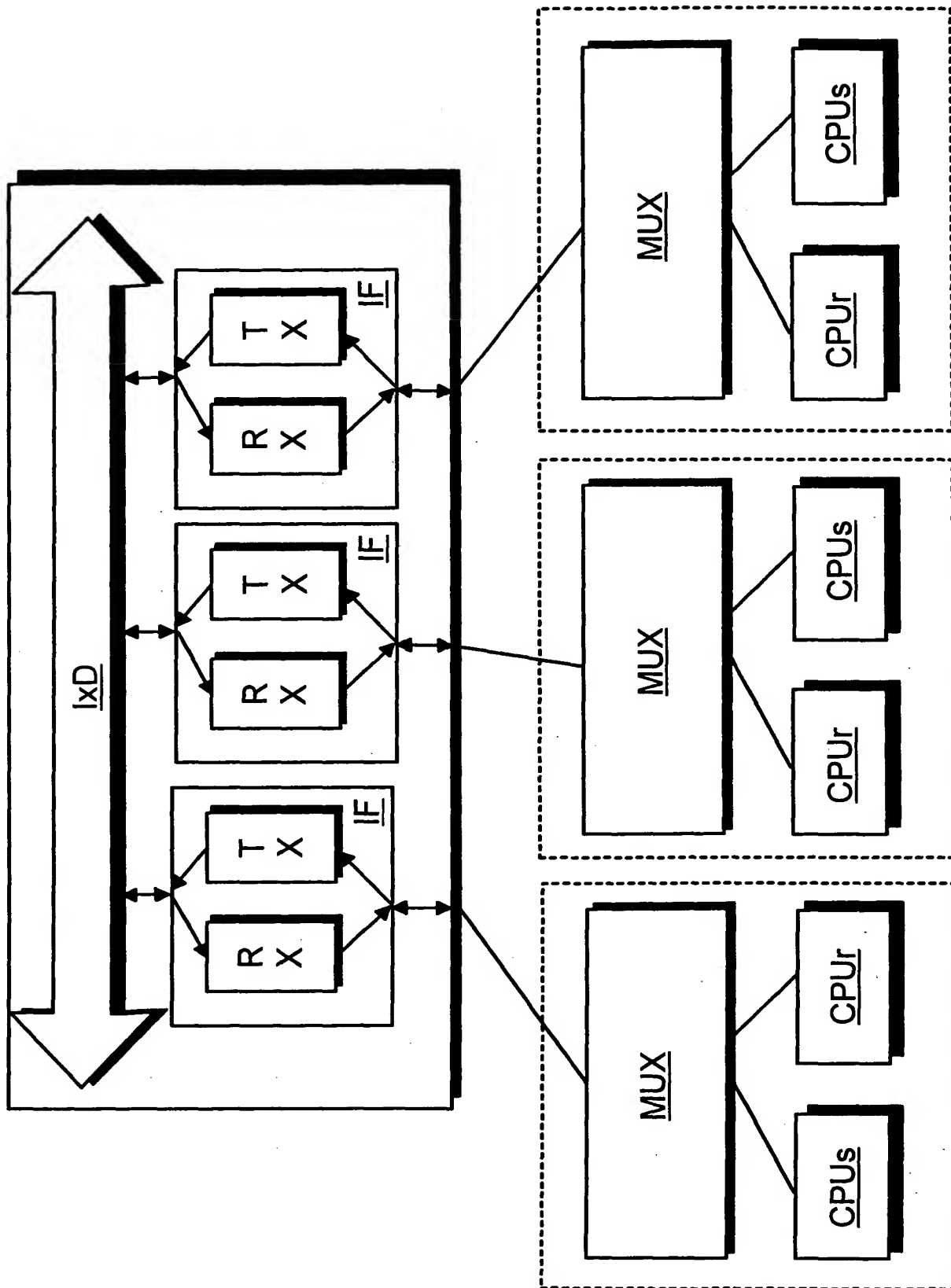


Fig. 3